q → τ and Entropic Tunneling: A Minimal-Corpus Route to Euler's Identity

Douglas C. Youvan

doug@youvan.com

www.youvan.ai

October 21, 2025

What is the smallest auditable body of facts from which one can rigorously derive Euler's identity, $e^{i\pi} + 1 = 0$? We pose this as a q → τ problem: start with a sparse internal model q (a minimal corpus of series axioms and complex arithmetic) and reach τ (the fully coherent target) along a low-action path. Here action is the prefix-free (MDL) description length of a proof script, which coincides with probability under a step model q_θ: $I(P) \approx -\log\ q_\theta(P)$. In expectation, action equals cross-entropy, $\mathbb{E}[I] = H(T) + \mathrm{KL}(T \parallel q_\theta)$, for a τ-consistent reference T. Thus height is the KL term—how far the learner's dynamics are from τ—and local curvature is governed by Fisher information (the quadratic expansion of KL). Width is distinct: representational multiplicity from commuting, symmetry-preserving moves that leave the intermediate normal form unchanged. We give a pen-and-paper derivation of Euler's identity inside a tightly audited series corpus (series definitions, $i^2 = -1$, Cauchy product, minimal π), and we prove a closed-form Entropic Tunneling law: if a segment admits S commuting rearrangements with identical continuation, then the minimal action drops by $\sim \log S$ without adding facts (a pure width/entropy bonus, orthogonal to KL height). Beauty, in this view, is geometry: invariance widens corridors, straightening the path to τ. We close with audit rules, negative controls, and extensions to ODE and rotation-group routes.

Keywords: q → τ, Entropic Tunneling, Euler's identity, minimal corpus, KL divergence, cross-entropy, Fisher information, information geometry, commuting symmetries, MDL, prefix-free coding, Cauchy product, series method, action minimization, proof audit.

A collaboration with GPT-5. CC4.0.

Outline

Premise and Results
1.1 Statement of goals (q → τ)
1.2 Main theorems: Euler inside a minimal series corpus; log S action drop

Minimal Corpus and Guardrails (q)
2.1 Primitives and axioms for exp, sin, cos (power series)
2.2 Algebra for absolutely convergent series; $i^2 = -1$
2.3 Definition of π; forbidden facts; audit protocol
2.4 Height and local geometry: KL and Fisher (definitions and role)

Core Lemmas from the Corpus
3.1 Even/odd split: $\exp(ix) = \cos x + i\sin x$
3.2 Termwise calculus: $\sin' = \cos$, $\cos' = -\sin$
3.3 $\sin^2 + \cos^2 = 1$; existence of a least positive zero of $\sin$
3.4 Monotonicity on $(0, \pi)$and $\cos(\pi) = -1$

Theorem (τ): Euler's Identity
4.1 Proof that $\exp(i\pi) = -1$within the corpus
4.2 Audit trail: line-by-line rule citations

Entropic Tunneling Theorem
5.1 Proof scripts as prefix-free codes (MDL action)
5.2 Commuting multiplicity S and equivalence classes
5.3 Theorem: $I^* \leq I_0 - \log S + O(1)$ (no new facts)
5.4 Probabilistic restatement via exchangeability (link to KL)

Negative Controls and Robustness
6.1 Smuggling tests (remove a rule; proof fails or inflates action)
6.2 Diffusion traps (cheap non-commuting moves flatten the gain)
6.3 Corpus minimality and MDL of the rule set

Extensions and Comparisons
7.1 ODE route within its own minimal corpus
7.2 Rotation-group route and complex-plane identification
7.3 Cross-route invariants; curriculum and ATP implications

Discussion and Outlook
8.1 Beauty as width: symmetry, invariance, and straightness of path
8.2 Libraries of minimal corpora for landmark identities
8.3 Program for larger τ: from identities to theorems

References

1. Premise and Results

1.1 Statement of goals (q → τ)

Objective. Establish a rigorously audited derivation of

$$\tau: e^{i\pi} + 1 = 0$$

from a minimal series corpus q(definitions of $\exp$, $\sin$, $\cos$ by Maclaurin series; $i^2 = -1$; Cauchy product; termwise calculus; basic field laws; minimal π as the least positive zero of $\sin$). In parallel, formalize a quantitative notion of Entropic Tunneling: holding q fixed (no new facts), we show that representational width—the number S of symmetry-preserving, commuting rearrangements in a proof segment—reduces the least action needed to reach τ.

Action, height, and width. A proof script is a finite sequence of licensed steps, encoded by a prefix-free code; its action is the code length I. With a step model q_θ $(\cdot|$ state$)$ and a τ-consistent reference distribution T,

$$\mathbb{E}_T[I] \;=\; H(T) + \mathrm{KL}(T \parallel q_\theta).$$

Thus height is the KL term (how far the prover's dynamics are from τ), and its local curvature is governed by Fisher information (the quadratic expansion of KL). Width is distinct: it is the entropy of commuting equivalence inside a corridor—reorderings that leave the intermediate normal form and continuation unchanged.

Goals.

1. (q ⊢ τ) Prove $\exp(i\pi) = -1$i nside q with a line-by-line audit.
2. (Width law) Prove a coding-theoretic bound showing an explicit $-\log S$ drop in minimal action from commuting multiplicity alone, orthogonal to the KL "height".

1.2 Main theorems: Euler inside a minimal series corpus; $\log S$action drop

Theorem A (Euler inside the minimal series corpus).
Within q as specified above,

$$\exp(i\pi) + 1 = 0.$$

Route. Using only series axioms and series algebra: (i) even/odd split of $\exp(ix)$ yields $\exp(ix) = \cos x + i\sin x$; (ii) termwise calculus gives

$$\sin' = \cos, \qquad \cos' = -\sin;$$

(iii) hence $\sin^2 + \cos^2 = 1$; (iv) with π the least positive zero of $\sin$, we have $\sin(\pi) = 0$ and strict decrease of $\cos$ on $(0, \pi)$, so $\cos(\pi) = -1$; evaluating (i) at $x = \pi$ gives $\exp(i\pi) = -1$. No ODE, rotation, or pre-assumed Euler facts are used.

Theorem B (Entropic Tunneling: $-\log S$ drop without new facts).
Fix q and a prefix-free encoding of scripts. Let Σ be a contiguous block of commuting, symmetry-preserving rewrites whose permutations (of size S) all reach the same checkpoint normal form and identical continuation. Let I_0 be the least action when Σ is spelled out in a single fixed order. Then

$$I^* \leq I_0 - \log_2 S + O(1).$$

Idea. Replace the explicit enumeration of Σ by a constant-length macro plus, at most, a uniform index in $\{1, \ldots, S\}$. By Kraft–McMillan, choosing among S exchangeable permutations costs $\log_2 S$ bits; eliminating that choice yields the stated saving. This improves action while holding the corpus (facts) and the step model q_θ fixed; it is a width/entropy bonus distinct from the KL "height".

Consequences.

- The Euler derivation is auditable and minimal within the series corpus.
- Under natural MDL/ exchangeability assumptions, symmetry widens corridors, straightening the path to $\tau$by an explicit $-\log S$ reduction in least action, independent of any change to KL/height.

2. Minimal Corpus and Guardrails (q)

2.1 Primitives and axioms for exp, sin, cos (power series)

Domain and symbols.

- Number systems: $\mathbb{R}$, $\mathbb{C}$.
- Imaginary unit $i \in \mathbb{C}$ with $i^2 = -1$.
- Variables: $x \in \mathbb{R}$, $z \in \mathbb{C}$.

- Functions: $\exp : \mathbb{C} \to \mathbb{C}$, $\sin$, $\cos : \mathbb{R} \to \mathbb{R}$.

Series axioms (definitions).

- (A1) Exponential (entire):

$$\exp(z) := \sum_{n\geq 0} \frac{z^n}{n!} (\forall z \in \mathbb{C}).$$

- (A2) Sine (real-analytic):

$$\sin(x) := \sum_{n\geq 0} (-1)^n \frac{x^{2n+1}}{(2n+1)!} (\forall x \in \mathbb{R}).$$

- (A3) Cosine (real-analytic):

$$\cos(x) := \sum_{n\geq 0} (-1)^n \frac{x^{2n}}{(2n)!} (\forall x \in \mathbb{R}).$$

Immediate consequences.

- (A4) Convergence/regularity: The series in (A1)–(A3) are absolutely convergent on their domains; hence exp is entire, $\sin$, cos are real-analytic.
- (A5) Base values and parity: $\exp(0) = 1$, $\sin(0) = 0$, $\cos(0) = 1$; sin odd, cos even.

These clauses *define* the functions purely by their Maclaurin series—no ODE, rotation, or geometric facts are presumed.

2.2 Algebra for absolutely convergent series; $i^2 = -1$

Complex field laws.

- (C1) $(\mathbb{C}, +, \cdot)$ is a field; $i^2 = -1$. For $a + ib, c + id \in \mathbb{C}$, $(a + ib)(c + id) = (ac - bd) + i(ad + bc)$.

Series algebra (valid under absolute convergence / analyticity).

- (S1) Termwise addition: if $\sum a_n$ and $\sum b_n$ converge absolutely, then $\sum(a_n + b_n)$ represents $f + g$.

- (S2) Cauchy product: for power series
  $f(x) = \sum_{n\geq 0} \quad a_n x^n$, $g(x) = \sum_{n\geq 0} \quad b_n x^n$,

$$f(x)g(x) = \sum_{n\geq 0} \quad (\sum_{k=0}^{n} \quad a_k b_{n-k}) x^n$$

on the common domain of absolute convergence (and globally for entire series).

- (S3) Termwise differentiation/integration: within absolute convergence,

$$\frac{d}{dx}(\sum_{n\geq 0} \quad a_n x^n) = \sum_{n\geq 1} \quad n a_n x^{n-1}, \int_0^x \quad (\sum_{n\geq 0} \quad a_n t^n)\, dt = \sum_{n\geq 0} \quad a_n \frac{x^{n+1}}{n+1}.$$

- (S4) Substitution by linear arguments: substitution $x \mapsto \alpha x + \beta$ with $\alpha, \beta \in \mathbb{C}$ commutes with the series operations when absolute convergence holds (e.g., $\exp(ix)$ via $z = ix$).

Permitted local rewrites. Finite regroupings (even/odd split), reindexing bijections, pulling out constants, distributing products over sums, and the applications of (S1)–(S4), (C1).

2.3 Definition of π; forbidden facts; audit protocol

Minimal π.
Let

$$\pi := \min\{x > 0 : \mathrm{s} \qquad\qquad \mathrm{in}(x) = 0\}.$$

Existence of a positive zero follows within the series framework by continuity (absolute convergence) and a concrete alternating-series sign change (e.g., at $x = 4$); least-positivity uses completeness and continuity (details provided in the lemmas).

Forbidden facts (to prevent smuggling).

- (F1) No prior use of Euler's formula $\exp(ix) = \cos x + i \sin x$.
- (F2) No ODE solution/uniqueness (e.g., "$f' = if \Rightarrow f = C\exp(ix)$").
- (F3) No rotation-group or matrix-exponential facts.
- (F4) No trigonometric identities beyond those derived from (A2)–(S3).
- (F5) No geometric definition of π(circumference/diameter), area/arc-length facts.

- (F6) No complex-integration theorems (Cauchy, residues), no Fourier representations.
- (F7) No external limit theorems beyond what (S1)–(S3) already guarantee under absolute convergence.

Audit protocol.

- (AP1) Line citation: each proof step cites exactly one of (A1)–(A5), (S1)–(S4), (C1) or elementary real-analysis facts used only for the existence/least-positivity of π.
- (AP2) Checkpoints: designate normal forms (e.g., "even/odd split completed"). Between checkpoints only commuting, symmetry-preserving rewrites are used—this isolates width S.
- (AP3) Forbidden-screen: for each lemma, list tempting external facts and mark them unused.
- (AP4) Ablation/minimality: remove one rule and retry the derivation; record Fail or Inflate (least action I^*strictly increases).
- (AP5) Width ledger: at each commuting block, record the set size $S$of orderings that reach the same checkpoint and identical continuation.

2.4 Height and local geometry: KL and Fisher (definitions and role)

Step model and action.
Equip the prover with a step distribution $q_\theta(\cdot|\ \text{state})$. Encode scripts with a fixed prefix-free code; for a script P,

$$I(P) \ \approx \ -\log\, q_\theta(P).$$

Let T be any τ-consistent reference distribution over correct next steps.

Height = KL (expected action).
By Kraft–McMillan / MDL,

$$\mathbb{E}_T[I] \ = \ H(T) \ + \ \mathrm{KL}\,(T \parallel q_\theta).$$

Thus height is the $\mathrm{KL}(T \parallel q_\theta)$ term: it quantifies how far the prover's dynamics are from the τ-consistent generator T. Lower action corresponds to tightening KL.

Local geometry = Fisher.
For small parameter changes,

$$\mathrm{KL}\,(T \parallel q_{\theta+\mathrm{d}\theta}) \;=\; \frac{1}{2}\,\mathrm{d}\theta^{\top}\,\mathcal{I}(\theta)\,\mathrm{d}\theta \;+\; o(\parallel \mathrm{d}\theta \parallel^{2}),$$

where $\mathcal{I}(\theta)$ is the Fisher information. Fisher gives the local curvature of height; in our discrete proxy (used later for step-weighting), moves aligned with symmetry manifolds effectively reduce curvature contributions while leaving provability unchanged.

Width (orthogonal to KL).
Let Σ be a corridor of commuting, symmetry-preserving rewrites whose permutations (size S) reach the same checkpoint and continuation. Exchangeability implies the class probability multiplies by S; hence

$$I \mapsto I - \log\, S(\text{bits}),$$

a pure entropy bonus independent of changing q_θ. This is the Entropic Tunneling effect proved later: height (KL/Fisher) fixes what is hard; width ($\log S$) shortens the description of *how* we traverse the permitted corridor—without adding facts.

3. Core Lemmas from the Corpus

All steps below use only the corpus rules from §2: series definitions (A1–A3), absolute-convergence algebra (S1–S4), complex field laws with $i^2 = -1$(C1), base values/parity (A5), and elementary real analysis for continuity/IVT. No ODE uniqueness, no rotation facts, and no pre-assumed Euler formula are used.

3.1 Even/odd split: $\exp\,(ix) = \cos\,x + i\sin\,x$

Claim. For all real x,

$$\exp\,(ix) \;=\; \cos\,x \;+\; i\,\sin\,x.$$

Proof (series only). By (A1) and $i^2 = -1$,

$$\exp\,(ix) = \sum_{n\geq 0} \frac{(ix)^n}{n!} = \sum_{k\geq 0} \frac{(ix)^{2k}}{(2k)!} + \sum_{k\geq 0} \frac{(ix)^{2k+1}}{(2k+1)!}.$$

Using (C1), $(ix)^{2k} = (i^2)^k x^{2k} = (-1)^k x^{2k}$ and $(ix)^{2k+1} = i(i^2)^k x^{2k+1} = i(-1)^k x^{2k+1}$. Hence, by (A3)–(A2),

$$\exp(ix) = \sum_{k\geq 0} \frac{(-1)^k x^{2k}}{(2k)!} + i \sum_{k\geq 0} \frac{(-1)^k x^{2k+1}}{(2k+1)!} = \cos x + i \sin x.$$

□

3.2 Termwise calculus: $\sin' = \cos$, $\cos' = -\sin$

Claim. For all real x,

$$\frac{d}{dx}\sin x = \cos x, \frac{d}{dx}\cos x = -\sin x.$$

Proof (termwise differentiation). Differentiate the defining series using (S3):

$$\frac{d}{dx}\sin x = \sum_{n\geq 0} (-1)^n \frac{(2n+1)x^{2n}}{(2n+1)!} = \sum_{n\geq 0} (-1)^n \frac{x^{2n}}{(2n)!} = \cos x,$$
$$\frac{d}{dx}\cos x = \sum_{n\geq 1} (-1)^n \frac{(2n)x^{2n-1}}{(2n)!} = -\sum_{n\geq 0} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = -\sin x.$$

□

3.3 $\sin^2 + \cos^2 = 1$; existence of a least positive zero of sin

(i) Identity $\sin^2 x + \cos^2 x = 1$

Proof. Let $h(x) = \sin^2 x + \cos^2 x$. Using the product/chain rules inherited from (S3) and §3.2,

$$h'(x) = 2\sin x \cos x + 2\cos x(-\sin x) = 0.$$

Thus h is constant; by (A5), $h(0) = 0^2 + 1^2 = 1$. Hence $\sin^2 + \cos^2 = 1$. □

(ii) Existence of a positive zero and least positivity

Step 1 (positivity near 0). From (A2),

$$\sin\ x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots ,$$

so for sufficiently small $x > 0$ the leading term dominates and $\sin\ x > 0$. Absolute convergence (A4) gives continuity.

Step 2 (sign change at a concrete point). At $x = 4$,

$$\sin\ 4 = 4 - \frac{4^3}{3!} + \frac{4^5}{5!} - \frac{4^7}{7!} + \cdots$$

is an alternating series with eventually decreasing term magnitudes; the displayed partial sum is negative and the tail is bounded by the next term, so $\sin\ 4 < 0$.

Step 3 (first zero and least positivity). By IVT, some $x_1 \in (0,4)$ has $\sin\ x_1 = 0$. Let $Z = \{x > 0 : \sin x = 0\}$ and $\pi = \inf\ Z$. Nonemptiness and boundedness below give $\pi \in (0,4]$. Continuity yields $\sin \pi = \lim_{n} \sin x_n = 0$ for any $x_n \downarrow \pi$ from Z. Thus π is the least positive zero. □

3.4 Monotonicity on $(0, \pi)$ and $\cos\ (\pi) = -1$

Claim. On $(0, \pi)$ one has $\sin x > 0$ and therefore $\cos\ '(x) = -\sin\ x < 0$; hence cos is strictly decreasing on $(0, \pi)$. Consequently,

$$\cos\ (\pi) = -1.$$

Proof. By definition of π as the least positive zero, $\sin x > 0$ for $x \in (0, \pi)$. From §3.2, $\cos\ '(x) = -\sin x < 0$ on that interval, so cos is strictly decreasing there. From §3.3(i), $\cos^2(\pi) = 1$. Since cos decreases from $\cos(0) = 1$ down to its endpoint value at $x = \pi$, it cannot return to $+1$; therefore $\cos\ (\pi) = -1$. □

Summary of §3. The corpus alone yields:

- the series-only identity $\exp\ (ix) = \cos\ x + i\sin\ x$;
- the calculus relations $\sin\ ' = \cos$, $\cos\ ' = -\sin$;

- the Pythagorean identity $\sin^2 + \cos^2 = 1$ and a least positive zero π of sin ;
- strict decrease of cos on $(0, \pi)$ and $\cos(\pi) = -1$.

Evaluating $\exp(ix) = \cos x + i \sin x$ at $x = \pi$ gives $\exp(i\pi) = -1$, i.e., τ.

4. Theorem (τ): Euler's Identity

4.1 Proof that $\exp(i\pi) = -1$ within the corpus

Theorem (τ).

$$\exp(i\pi) + 1 = 0.$$

Proof (using only q).

1. Series decomposition. By the exponential series (A1) and $i^2 = -1$ (C1),

$$\exp(ix) = \sum_{n \geq 0} \frac{(ix)^n}{n!} = \sum_{k \geq 0} \frac{(ix)^{2k}}{(2k)!} + \sum_{k \geq 0} \frac{(ix)^{2k+1}}{(2k+1)!}.$$

Using $(ix)^{2k} = (i^2)^k x^{2k} = (-1)^k x^{2k}$ and $(ix)^{2k+1} = i(i^2)^k x^{2k+1} = i(-1)^k x^{2k+1}$ (C1),

$$\exp(ix) = \sum_{k \geq 0} \frac{(-1)^k x^{2k}}{(2k)!} + i \sum_{k \geq 0} \frac{(-1)^k x^{2k+1}}{(2k+1)!}.$$

By the defining series (A3)–(A2),

$$\exp(ix) = \cos x + i \sin x. \qquad (4.1)$$

2. Minimal π. Let π be the least positive zero of sin (Def. §2.3). From §3.3(ii), such π exists and satisfies $\sin(\pi) = 0$.
3. Endpoint cosine value. On $(0, \pi)$ we have $\sin x > 0$, hence by §3.2, $\cos'(x) = -\sin x < 0$: cos is strictly decreasing there. With $\sin^2 + \cos^2 = 1$ (§3.3(i)), it follows that $\cos(\pi) = -1$.
4. Evaluation at π. Substitute $x = \pi$ into (4.1) and use Steps 2–3:

$$\exp(i\pi) = \cos \pi + i \sin \pi = (-1) + i \cdot 0 = -1.$$

Therefore $\exp(i\pi) + 1 = 0$. □

4.2 Audit trail: line-by-line rule citations

Each transformation below cites exactly one allowed rule (from §2); no forbidden fact is used.

1. Introduce $E(x)$. Let $E(x) := \exp(ix)$.
   [notation only]

2. Expand exp .
   $E(x) = \sum_{n\geq 0} (ix)^n/n!$.
   Rules: (A1), substitution $z = ix$ via (S4).

3. Even/odd split of the sum.
   $\sum_{n\geq 0} (ix)^n/n! = \sum_{k\geq 0} (ix)^{2k}/(2k)! + \sum_{k\geq 0} (ix)^{2k+1}/(2k+1)!$.
   Rules: (S1) regrouping under absolute convergence; reindexing (bijective).

4. Reduce powers of i; factor i from odd part.
   $(ix)^{2k} = (i^2)^k x^{2k} = (-1)^k x^{2k}$;
   $(ix)^{2k+1} = i(i^2)^k x^{2k+1} = i(-1)^k x^{2k+1}$.
   Rules: (C1).

5. Identify cos and sin series.
   $\sum_{k\geq 0} (-1)^k x^{2k}/(2k)! = \cos x$,
   $\sum_{k\geq 0} (-1)^k x^{2k+1}/(2k+1)! = \sin x$.
   Rules: (A3), (A2).
   Checkpoint 1: $\exp(ix) = \cos x + i\sin x$.

6. Define π and record its property.
   $\pi = \min\{x > 0 : \sin x = 0\}$ and $\sin(\pi) = 0$.
   Rules: Def. §2.3; existence from §3.3(ii).

7. Differential relations.
   $\sin' = \cos$, $\cos' = -\sin$.
   Rules: (S3) applied to (A2)–(A3) (as in §3.2).

8. Monotonicity on $(0, \pi)$.
   Since $\sin x > 0$ on $(0, \pi)$ (least positivity), $\cos'(x) = -\sin x < 0$; thus cos strictly decreases.
   Rules: Step 7; elementary order on $\mathbb{R}$.

9. Pythagorean identity.
   $\sin^2 + \cos^2 = 1$ and hence $\cos^2(\pi) = 1$.
   Rules: derivative of $h = \sin^2 + \cos^2$ via (S3) and Step 7; evaluate at 0 using (A5).

10. Endpoint value.
    From monotonicity (Step 8) and $\cos^2(\pi) = 1$ (Step 9), conclude $\cos(\pi) = -1$.
    Rules: monotone-limit reasoning.

11. Evaluate Checkpoint 1 at $x = \pi$.
    $\exp(i\pi) = \cos\pi + i\sin\pi = (-1) + i \cdot 0 = -1$.
    Rules: substitution; Steps 6 & 10; (C1).

12. Conclude τ.
    $\exp(i\pi) + 1 = 0$.
    Rules: (C1).

Forbidden-screen. No step uses Euler's formula as an assumption, ODE uniqueness, rotation/matrix facts, non-series trig identities, or complex integration (F1–F7).

5. Entropic Tunneling Theorem

5.1 Proof scripts as prefix-free codes (MDL action)

Instruction alphabet. Let Abe the finite set of licensed local steps from §2—applications of (A1–A5), (S1–S4), (C1), plus finitely many administrative macros (e.g., "reindex", "collect real/imag parts") each definable as a finite sequence over A itself.

Scripts and decoding. A *script* is a finite word $w = a_1 a_2 \cdots a_m \in A^*$ that, when executed from the initial statement ("start with $\exp(ix)$"), yields a terminal statement. A script is *successful for* τ if its terminal statement is exactly $\exp(i\pi) + 1 = 0$.

Prefix-free code and action. Fix any prefix-free code $C: A \to \{0,1\}^+$, extended to words by concatenation. Define the action of $w$by

$$I(w) := |\, C(w) \,| \text{ (bits).}$$

The least action for τ is

$$I^* := \inf\{I(w) : w \text{ is successful for } \tau\}.$$

By standard MDL/Kraft–McMillan facts, changing the underlying prefix-free code perturbs actions by at most an additive constant $O(1)$; asymptotic and $\log S$ statements are invariant up to $O(1)$.

Normalization checkpoints. We fix canonical *checkpoints* (e.g., "even/odd split completed"). Between two consecutive checkpoints we only allow local rewrites; this isolates the corridor where commuting multiplicity will be counted.

5.2 Commuting multiplicity $S$and equivalence classes

Commuting block Σ. Consider a contiguous block Σ inside a successful script such that:

1. (Locality) Every instruction in Σ is a licensed local rewrite from §2.
2. (Confluence to a checkpoint) Regardless of order, maximal application of these rewrites produces the same intermediate normal form $NF(\Sigma)$ and the same continuation thereafter.
3. (Pairwise commutation) Whenever two instructions act on disjoint or commuting subexpressions, their effects commute as rewrite relations: $u \circ v = v \circ u$.

Equivalence of scripts on Σ. Two successful scripts are *equivalent on* Σ if they are identical outside Σ and their substrings on Σ are permutations of the same multiset of instructions (hence both reach $NF(\Sigma)$ and then continue identically). Let E_Σ be the set of such scripts, with multiplicity

$$S := | E_\Sigma |.$$

Intuitively, S is the *width*—the number of symmetry-preserving orderings that leave semantics (checkpoint + continuation) unchanged.

5.3 Theorem: $I^* \le I_0 - \log_2 S + O(1)$ (no new facts)

Setup. Let I_0 be the least action among successful scripts that traverse Σ by spelling out one fixed order of its steps (keeping the rest of the script fixed).

Macro encoding of Σ. Introduce a macro symbol $M_\Sigma \in A$ (or a short macro word over A) that denotes: "apply the Σ-block rewrites until the checkpoint $NF(\Sigma)$ is reached." Because M_Σ expands to a finite, corpus-legal sequence, we can assign it a fixed codeword of length

$O(1)$ (independent of S). If one must *reproduce a particular ordering* within Σ (e.g., for deterministic replay), append a uniform binary index in $\{1, \ldots, S\}$.

Coding lemma (Kraft–McMillan). Choosing uniformly among S alternatives costs $\lceil \log_2 S \rceil$ bits in any prefix-free encoding; this is optimal up to $O(1)$.

Theorem 5.1 (Entropic Tunneling; MDL form).
For any commuting block Σ with multiplicity $S \geq 1$,

$$I^* \leq I_0 - \log_2 S + O(1),$$

where the constant is independent of S.

Proof. Take a script w_0 achieving I_0. Replace its explicit Σ-substring by M_Σ (constant code length) plus, at most, a $\lceil \log_2 S \rceil$-bit index selecting the desired ordering. Since every ordering in E_Σ yields the same $NF(\Sigma)$ and identical continuation, the replacement gives a successful script of action

$$I(w_0) - L_{\text{explicit}}(\Sigma) + (|\, C(M_\Sigma) \,| + \lceil \log_2 S \rceil).$$

Any explicit encoding must, at minimum, identify which of the S orderings is realized to remain prefix-free and reconstructible, so $L_{\text{explicit}}(\Sigma) \geq \lceil \log_2 S \rceil - O(1)$. Hence the new action is $\leq I_0 - \log_2 S + O(1)$. Minimizing over w_0 yields the bound for I^*. □

No new facts (height unchanged). The macro denotes a sequence already legal in the corpus. We alter only the *code* (representation), not the *rule set* or provable statements. Thus the "height" (what is provable) is fixed; the reduction is a width/entropy effect.

When the index is unnecessary. If the post-checkpoint continuation depends only on $NF(\Sigma)$ (as in our Euler derivation), we omit the index; then the reduction is $\geq \log_2 S - O(1)$.

5.4 Probabilistic restatement via exchangeability (link to KL)

Step model. Let $q_\theta(\cdot|\ \text{state})$ be a prover's step distribution. For a script w, define its probability $Q_\theta(w)$ as the product of step probabilities along the path.

Exchangeability on Σ. Suppose q_θ treats the commuting permutations in E_Σ exchangeably (equal probability given identical contexts). Then the class probability

$$Q_\theta(E_\Sigma) = \sum_{w \in E_\Sigma} Q_\theta(w) = S \cdot Q_\theta(w_0)$$

for any representative $w_0 \in E_\Sigma$. The corresponding action (negative log-probability) of the class drops by

$$-\log_2 Q_\theta(E_\Sigma) \;=\; -\log_2 Q_\theta(w_0) \;-\; \log_2 S.$$

Thus the width contributes a $-\log_2 S$ term—identical to the MDL bound.

Connection to KL height. If T is a τ-consistent reference over scripts, the expected action satisfies

$$\mathbb{E}_T[I] \;=\; H(T) + \mathrm{KL}(T \,\|\, Q_\theta).$$

Exchangeability multiplies the mass of each commuting class by S, subtracting $\log_2 S$ from $\mathbb{E}_T[I]$ without changing Q_θ outside Σ. Hence the tunneling gain is a pure width effect, orthogonal to KL "height": we compress *how* we traverse an already-valid corridor, we do not alter *what* can be proved or the step model's fit to T.

5. Entropic Tunneling Theorem

5.1 Proof scripts as prefix-free codes (MDL action)

Instruction set. Let A be the finite alphabet of licensed local steps from §2 (applications of (A1–A5), (S1–S4), (C1)) plus a finite set of administrative primitives (e.g., *reindex*, *collect real/imag*), each definable as a finite sequence over the same laws.

Scripts and success. A script is a finite word $w = a_1 \dots a_m \in A^{\setminus *}$. Executed from the initial statement ("start with $\exp(ix)$"), w is successful for τ if its terminal line is $\exp(i\pi) + 1 = 0$.

Prefix-free action. Fix any prefix-free code $C: A \to \{0,1\}^+$, extend by concatenation to words, and define

$$I(w) := |\, C(w) \,| \text{ (bits)}, \; I^* := \inf\{I(w) : w \text{ successful for } \tau\}.$$

Changing C perturbs all actions by at most an additive constant $O(1)$ (Kraft–McMillan), so any $\log S$ claim is invariant up to $O(1)$.

Checkpoints. Choose canonical checkpoints (e.g., "even/odd split completed"). Between two checkpoints, we allow only local rewrites; this isolates the corridor segments where commuting multiplicity will be measured.

5.2 Commuting multiplicity S and equivalence classes

Commuting block. Let Σ be a contiguous block inside some successful script such that:

1. (Locality) Each instruction in Σ is a licensed local rewrite from §2.
2. (Confluence to a checkpoint) All maximal applications of Σ's rewrites—regardless of order—reach the same intermediate normal form $NF(\Sigma)$ and then continue identically to τ.
3. (Pairwise commutation) For any two instructions u, v in Σ that act on disjoint/commuting subexpressions, $u \circ v = v \circ u$ as rewrite relations.

Equivalence. Two successful scripts w, w' are equivalent on Σ if they are identical outside Σ and their substrings on Σ are permutations of the same multiset of instructions (hence both reach $NF(\Sigma)$ and share the same tail). Let

$$E_\Sigma := \{\, w : w \text{ successful, equivalent on } \Sigma \,\}, S := |\, E_\Sigma \,|.$$

Intuition: S is the width—the number of symmetry-preserving orderings that do not alter semantics (checkpoint + continuation).

5.3 Theorem: $I^* \le I_0 - \log_2 S + O(1)$ (no new facts)

Setup. Let I_0 be the least action among successful scripts that traverse Σ by spelling out one fixed order of its steps (rest of the script fixed).

Macro for Σ. Introduce a macro symbol M_Σ that denotes: "apply the Σ-block rewrites until $NF(\Sigma)$ is reached." Because M_Σ expands to a finite, corpus-legal sequence, we can assign it a codeword of length $O(1)$, independent of S. If one must reproduce a particular ordering inside Σ, append a uniform index in $\{1, \dots, S\}$.

Coding lemma. Any prefix-free encoding that distinguishes S equiprobable alternatives uses $\lceil \log_2 S \rceil$ bits, optimal up to $O(1)$.

Theorem 5.1 (Entropic Tunneling; MDL form).
For any commuting block Σ with multiplicity $S \ge 1$,

$$I^* \;\le\; I_0 \;-\; \log_2 S \;+\; O(1),$$

where the $O(1)$ constant is independent of S.

Proof. Take a script w_0 achieving I_0. Replace the explicit Σ-substring by M_Σ (constant length) plus, if necessary, a $\lceil \log_2 S \rceil$-bit index. The new action is

$$I(w_0) - L_{\text{explicit}}(\Sigma) + (|\, C(M_\Sigma)\,| + \lceil \log_2 S \rceil).$$

Any explicit encoding must identify which of the S orderings occurs (for prefix-freeness and replay), so $L_{\text{explicit}}(\Sigma) \geq \lceil \log_2 S \rceil - O(1)$. Hence the action drops by at least $\log_2 S - O(1)$. Minimizing over w_0 yields the bound for I^*. □

No new facts. The macro abbreviates a sequence already legal under the corpus; the rule set and provability are unchanged. The gain is a representation (coding) effect—pure width—not a change of "height."

Order-free tails. If the continuation depends only on $NF(\Sigma)$ (as in our Euler derivation), no index is needed, and the drop is $\geq \log_2 S - O(1)$.

5.4 Probabilistic restatement via exchangeability (link to KL)

Step model. Let $Q_\theta(w)$ be the script probability induced by a step model $q_\theta(\cdot \mid \text{state})$ (product of stepwise probabilities).

Exchangeability on Σ. If q_θ assigns equal weight to each permutation in E_Σ (they are indistinguishable given the same context), then

$$Q_\theta(E_\Sigma) \;=\; \sum_{w \in E_\Sigma} Q_\theta(w) \;=\; S \cdot Q_\theta(w_0),$$

for any representative $w_0 \in E_\Sigma$. The action of the class therefore satisfies

$$-\log_2 Q_\theta(E_\Sigma) \;=\; -\log_2 Q_\theta(w_0) \;-\; \log_2 S.$$

Thus width contributes exactly the $-\log_2 S$ bonus, matching the MDL bound.

Link to KL (height). For a τ-consistent reference distribution T over scripts,

$$\mathbb{E}_T[I] \;=\; H(T) + \mathrm{KL}(T \,\|\, Q_\theta).$$

Exchangeability multiplies probability mass of each commuting class by S, subtracting $\log_2 S$ from $\mathbb{E}_T[I]$ without changing Q_θ outside Σ. Hence the tunneling gain is orthogonal to

height: it compresses *how* we traverse a fixed corridor, not *what* is provable nor the KL fit of Q_θ to T.

6. Negative Controls and Robustness

6.1 Smuggling tests (remove a rule; proof fails or inflates action)

Purpose. Verify that τ is reached strictly within the stated corpus and that the $-\log S$ gain reflects genuine width, not hidden height (extra facts).

Protocol (single-rule ablation).

- For each rule $R \in \{(A1) \ldots (A5), (S1) \ldots (S4), (C1)\}$:
    1. Form $q \setminus \{R\}$.
    2. Attempt the derivation (keeping the route fixed and the same checkpoints).
    3. Record outcome: Fail (τ unreachable) or Inflate (τ reachable but least action I^* strictly larger).

Expected outcomes (series route).

- Remove (A1) (exp series): cannot start even/odd split ⇒ Fail.
- Remove (A2) or (A3) (sin/cos series): cannot identify even/odd parts ⇒ Fail.
- Remove (S3) (termwise differentiation): lose $\sin' = \cos$, $\cos' = -\sin$; can't certify monotonicity ⇒ typically Fail; any workaround (purely combinatorial bound at π) sharply Inflates I^*.
- Remove (S1) (regrouping): even/odd isolation breaks; explicit Cauchy products explode step count ⇒ Inflate.
- Remove (C1) (complex algebra, $i^2 = -1$): no reduction of $(ix)^n$⇒ Fail.

Two-rule ablation sanity checks.

- Removing $(S1)$ and $(S2)$ concurrently eliminates both regrouping and Cauchy product; τ must Fail.
- Removing $(A2)$ but adding any forbidden shortcut (e.g., ODE uniqueness) would reach τ; the audit flags this as smuggling.

Metric. Log the pair $(\Delta L_{\text{rules}}, \Delta I^*)$, where ΔL_{rules} is the saved rule-description bits, ΔI^* the increase in least action. A rule is non-redundant if $\Delta I^* > -\Delta L_{\text{rules}}$.

6.2 Diffusion traps (cheap non-commuting moves flatten the gain)

Risk. Adding many *cheap but non-commuting* rewrites increases apparent branching without forming a single equivalence class; the resulting "width" is noise, not symmetry. This can flatten or reverse the $-\log S$ slope.

Trap constructors (do *not* count toward S).

- Vacuous toggles: insert/remove 0 or multiply/divide by 1 in ways that interact non-commutatively with neighboring steps (e.g., $x \mapsto x + 0 \cdot x$ before/after collection).
- Ambiguous reindexing: non-bijective or scope-entangling reindex rules (break confluence to the checkpoint).
- Multiple non-canonical "normal forms": parallel pretty-printers whose outputs are not provably identical under the corpus.

Effect on action.
Let S_{true} be the count of commuting, symmetry-preserving permutations (as in §5.2) and M the number of added non-commuting detours. The MDL-optimal script must specify disambiguations for the detours; typically

$$I^*_{\text{trap}} \;\geq\; I^*_{\text{clean}} \;-\; \log_2 S_{\text{true}} \;+\; c \log_2 M,$$

with $c > 0$ depending on how many bits are needed to fix order/scope. Hence diffusion increases action; it does not realize the $-\log S$ bonus.

Guardrail for counting S.
A permutation counts toward Sonly if:

1. rewrites commute as relations on expressions;
2. they converge to the same checkpoint normal form;
3. the tail (continuation) is identical.
   Else, the move is excluded from $S$and treated as diffusion.

6.3 Corpus minimality and MDL of the rule set

Two-part MDL.
We score a proof system + proof by

$$\text{MDL}_{\text{total}}(\tau) \;=\; L_{\text{rules}}(q) \;+\; I^*(\tau \mid q),$$

where $L_{\text{rules}}(q)$ is a fixed prefix-free description of the corpus (axioms, permitted rewrites, and any named macros), and I^* is least action for τ.

Pareto-minimality (definition).
A corpus q is Pareto-minimal for τ if no strict subcorpus $q' \subset q$ satisfies

$$L_{\text{rules}}(q') + I'^*(\tau \mid q') \;<\; L_{\text{rules}}(q) + I^*(\tau \mid q).$$

Practical tests.

1. Single-rule MDL audit. For each rule R,

$$\Delta_{\text{MDL}}(R) \;=\; \underbrace{-\Delta L_{\text{rules}}(R)}_{\text{bits saved by removing } R} \;+\; \underbrace{\Delta I^*(R)}_{\text{extra proof bits without } R}.$$

Keep R iff $\Delta_{\text{MDL}}(R) > 0$.

2) Macro amortization. When introducing a macro M_Σ (used in §5), charge it once in L_{rules}. Over multiple theorems $\{\tau_j\}$, the amortized cost per theorem is $\frac{L_{\text{rules}}(M_\Sigma)}{N}$ for N uses; as N grows, the $-\log S$ savings dominates.

3) Cross-route leakage. If adding ODE or rotation facts shortens I^* on the series route but increases L_{rules} by more than the savings, it fails the MDL test (not minimal for this paper). Keep routes separate unless comparing them explicitly.

Robustness summary.

- Smuggling tests show essential rules are truly needed (or action-reducing).
- Diffusion controls ensure only *true* commuting symmetry contributes to S.
- MDL accounting enforces that our corpus is near-minimal: any removal breaks τ or inflates I^*; any addition that is not symmetry-preserving fails to reduce $\text{MDL}_{\text{total}}$.

These controls ensure that the observed $-\log S$ reduction is a principled width effect within a minimal, auditable q, not an artifact of hidden height or noisy search freedom.

7. Extensions and Comparisons

7.1 ODE route within its own minimal corpus

Minimal corpus q_{ode}.

- **Primitives:** $\mathbb{R}, \mathbb{C}$; $i^2 = -1$.
- **Axioms/definitions:**
  (O1) $\exp'(x) = \exp(x),\ \exp(0) = 1.$
  (O2) $\sin, \cos$ are the unique real-analytic solutions of

$$\sin'' = -\sin, \quad \sin(0) = 0,\ \sin'(0) = 1; \qquad \cos'' = -\cos, \quad \cos(0) = 1,\ \cos'(0) = 0.$$

  (O3) Linearity/product rule for derivatives; substitution/evaluation rules.
  (O4) Complex field laws (C1).

Derivation of the checkpoint and τ.
Let $f(x) = \cos x + i \sin x$. By (O2)–(O3), $f'(x) = -\sin x + i \cos x = i(\cos x + i \sin x) = i\, f(x)$ and $f(0) = 1$. Uniqueness for the first-order linear IVP implies $f(x) = \exp(ix)$ given (O1). Define π as the least positive zero of $\sin$ (continuity follows from (O2)). Then $\sin(\pi) = 0$; on $(0, \pi)$, $\sin > 0 \Rightarrow \cos' = -\sin < 0 \Rightarrow \cos(\pi) = -1$. Evaluating f at π yields $\exp(i\pi) = -1$ (τ).

Entropic tunneling in q_{ode}.
The commuting block Σ_{ode} arises from linear operations that commute: differentiation vs. substitution inside linear contexts, collecting real/imag parts, and reordering applications of linearity and evaluation before the uniqueness step. These permutations lead to the same checkpoint "$f'(x) = if(x), f(0) = 1$". If S_{ode} is the number of such permutations with identical continuation, then by the MDL argument,

$$I^*_{\mathrm{ode}} \leq I_{0,\mathrm{ode}} - \log_2 S_{\mathrm{ode}} + O(1).$$

No new facts are added; the reduction is purely representational width.

Forbidden (for audit): no use of power-series definitions or rotation matrices; the only uniqueness principle is that for linear ODEs with analytic data (bundled with (O2) as part of q_{ode}).

7.2 Rotation-group route and complex-plane identification

Minimal corpus q_{rot}.

- **Primitives:** vectors in $\mathbb{R}^2$, matrices, and the identification $(x, y) \leftrightarrow x + iy \in \mathbb{C}$.
- **Axioms/definitions:**
  (R1) Rotation matrices $R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ satisfy
  $R(\theta)R(\phi) = R(\theta + \phi),\ R(0) = I,\ \theta \mapsto R(\theta)$ continuous.
  (R2) Generator $J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, with matrix exponential
  $\exp(\theta J) = \sum_{n \geq 0} (\theta J)^n / n!$ (absolutely convergent).
  (R3) Identify multiplication by i on $\mathbb{C}$ with the linear map J on $\mathbb{R}^2$.

Derivation of the checkpoint and τ.
From (R2), $\exp(\theta J)$ is a one-parameter subgroup with derivative at 0 equal to J, hence coincides with $R(\theta)$ by the group law/uniqueness of one-parameter subgroups in $\mathrm{SO}(2)$ (this is a minimal group-theoretic fact included in q_{rot}). The half-turn is $R(\pi) = -I$. Under (R3), $\exp(\pi J)$ corresponds to $\exp(i\pi)$, and $-I$ corresponds to multiplication by -1. Therefore $\exp(i\pi) = -1$ (τ).

Entropic tunneling in q_{rot}.
The commuting block Σ_{rot} consists of group symmetries: rebracketing products (associativity), partitioning angles $\theta = \theta_1 + \cdots + \theta_k$ in different orders, and passing through the homomorphism $\exp(\theta J)$–$R(\theta)$ in either direction when the generator commutes (here J is fixed). All such permutations reach the checkpoint "half-turn $\Rightarrow -I$". If S_{rot} is the number of such permutations with identical continuation, then

$$I^*_{\mathrm{rot}} \leq I_{0,\mathrm{rot}} - \log_2 S_{\mathrm{rot}} + O(1).$$

Again, no new facts are added; only a coding gain from width.

Forbidden (for audit): no power-series or ODE uniqueness; trigonometric addition rules may be used only insofar as they are consequences of (R1).

7.3 Cross-route invariants; curriculum and ATP implications

Cross-route invariants.

- Common checkpoint (saddle-touch).
  - Series: $\exp(ix) \to \cos x + i\sin x$.
  - ODE: $f'(x) = if(x), f(0) = 1$.

- Rotation: $\exp(\theta J) = R(\theta)$.
  These express the same invariant: the one-parameter subgroup generated by i acts as a rotation.

- Width-only savings. In all routes, commuting permutations that preserve the checkpoint form equivalence classes of size S, giving the same $-\log S$ action drop under MDL/exchangeability.
- Height stability. None of the width manipulations changes provability or the step model; they only compress *how* a valid corridor is traversed.

Curriculum design.

- Scaffolded representations. Teach series → ODE → rotation as three views of the same invariant. Students repeatedly encounter the checkpoint, learning to recognize and exploit commuting freedom.
- Macro thinking. Encourage packaging commuting micro-steps as macros (even/odd regroupings; linear Leibniz blocks; group rebracketings). This mirrors the MDL gain and reduces cognitive load.
- Minimal πchoices. Keep πminimal in each route (first positive zero; half-turn) to avoid smuggling geometric facts.

Automated theorem proving (ATP).

- Symmetry-aware search. Detect commuting subgraphs in the rewrite DAG; canonicalize to checkpoints; score scripts with an entropy bonus $\approx \log S$.
- Route-switch proposals. Allow nonlocal switches (series ↔ODE ↔rotation) *only at checkpoints* to preserve invariants and expose larger S.
- Two-part MDL objective. Minimize $L_{\text{rules}} + I^*$: penalize rule-set bloat, reward compressed proofs; reject diffusion traps (non-commuting noise).
- Replica ladders. Run replicas with increasingly permissive commuting libraries to widen true corridors; select scripts whose action falls without invoking forbidden facts.

Takeaway. The series, ODE, and rotation routes are representation-equivalent tunnels to the same τ. Their difference is not what is provable but how wide the corridor is. Designing pedagogy and ATP systems to maximize true width—while keeping the corpus minimal and auditable—systematically straightens the path to elegant proofs.

8. Discussion and Outlook

8.1 Beauty as width: symmetry, invariance, and straightness of path

In this framework, beauty becomes operational: a proof is beautiful when true width—the size S of symmetry-preserving, commuting equivalence classes—makes the route to τ effectively straight. Height (difficulty) is encoded by expected action $H + \mathrm{KL}$, while width contributes a route-independent entropy bonus $-\log S$ that shortens descriptions without changing what is provable. Three linked ideas capture the phenomenon:

- Symmetry. Permutations of micro-moves (even/odd regroupings, linear Leibniz blocks, group rebracketings) commute and preserve a fixed checkpoint normal form. Their count is S.
- Invariance. Each route's checkpoint expresses the same invariant ("the 1-parameter subgroup generated by i rotates the plane"), so different representations share the same saddle-touch.
- Straightness. Canonicalizing commuting families into macros collapses detours; the MDL-optimal script aligns with symmetry, yielding visibly shorter, cleaner arguments.

Thus elegance is compression by invariance: a small corpus, few essential steps, and a principled $-\log S$ reduction from symmetry—not rhetorical brevity, but structural economy.

8.2 Libraries of minimal corpora for landmark identities

The Euler case suggests a scalable artifact: a library of minimal corpora ("identity cards") for widely taught results, each card containing:

1. Corpus q: primitives, axioms, allowed rewrites;
2. Forbidden set: smuggling guards;
3. Checkpoints: canonical normal forms;
4. $q \vdash \tau$ proof: line-by-line with citations;
5. Width ledger: enumerated commuting blocks and S;
6. MDL audit: $L_{\text{rules}} + I^*$ plus ablation outcomes.

Illustrative entries:

- Trig addition laws (series vs. rotations): width from associativity and parity; checkpoint "angle addition equals series/product composition."

- Binomial theorem (formal vs. analytic): width from term permutations; S~ multinomial symmetries.
- Fourier kernel symmetries (shift/modulation/parity): width from commuting phase operators; canonical frequency-domain checkpoints.
- Gamma identities (reflection/duplication): width from commuting product rearrangements and functional equations (within an audited corpus).

Such a library supports pedagogy (route selection by width and student background) and benchmarking for ATP systems (shared corpora, shared checkpoints, comparable I^* and S statistics).

8.3 Program for larger τ: from identities to theorems

Scaling beyond identities requires structure without sacrificing minimality.

- Layered checkpoints. Factor long arguments into tiers of checkpoints (lemmas as normal forms). If blocks Σ_j are locally commuting and independently canonical, the savings accumulate:

$$I^* \;\le\; I_0 \;-\; \sum_j \;\; \log\, S_j \;+\; O(1).$$

- Reusable macros. Promote recurring commuting blocks (series regrouping, linear ODE reduction, group factorizations) to named macros charged once in L_{rules}. With reuse across many τ, amortized cost per theorem dwindles while $-\log\, S$ gains persist.
- Route interchanges. Permit nonlocal switches (series ↔ ODE ↔ rotation) only at checkpoints preserving invariants; switches can unlock larger S in the next segment without enlarging the corpus.
- Complexity control (MDL). Enforce Pareto-minimality: add a rule only if it reduces $L_{\text{rules}} + I^*$. Reject diffusion moves that expand branching but fail the commuting/normal-form tests.
- Machine-checkable audits. For sizable τ, annotate scripts with rule IDs, checkpoint hashes, and an explicit Sledger; this enables reproducible action counts, ablation tests, and apples-to-apples comparisons across routes and systems.
- Targets. Natural next steps include uniqueness of power-series expansions, low-dimensional spectral theorems (where commuting normal forms are rich), and

classification of plane isometries—each with clearly identifiable checkpoints and commuting families.

Outlook. The $q \to \tau$ methodology reframes proof design as tunnel design: identify invariants, fix checkpoints, maximize true commuting width, and measure success by the drop in action—not by folklore difficulty. For teaching, this yields structured clarity; for ATP, a principled objective—beauty as width—that is both measurable and optimizable.

References

1. Youvan, D. C. (October 2025). *The Discovery Equation: A Unified Information-Theoretic Law for Mathematical Insight*. ResearchGate. https://www.researchgate.net/publication/396514065_The_Discovery_Equation_A_Unified_Information-Theoretic_Law_for_Mathematical_Insight
2. Cover, T. M., & Thomas, J. A. (2006). *Elements of Information Theory* (2nd ed.). Wiley.
3. Rissanen, J. (2007). *Information and Complexity in Statistical Modeling*. Springer. *(Foundational MDL/description-length framework.)*
4. Ahlfors, L. V. (1979). *Complex Analysis* (3rd ed.). McGraw–Hill. *(Classical treatment of power series and elementary complex functions.)*
5. Rudin, W. (1987). *Real and Complex Analysis* (3rd ed.). McGraw–Hill. *(Background on series, analyticity, and basic real analysis tools.)*
6. MacKay, D. J. C. (2003). *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press. *(Coding, prefix-free descriptions, and probabilistic modeling.)*